Learning from power system data stream

Mauro Escobar and Daniel Bienstock
Industrial Engineering and Operations Research
Columbia University, NY, USA
{me2533,dano}@columbia.edu

Michael Chertkov
Los Alamos National Laboratory, NM, USA and
Program in Applied Mathematics
University of Arizona, Tucson, AZ, USA
chertkov@math.arizona.edu

***Abstract*—Assuming access to synchronized stream of Phasor Measurement Unit (PMU) data over a significant portion of a power system interconnect, say controlled by an Independent System Operator (ISO), what can you extract about past, current and future state of the system? We have focused on answering these practical questions pragmatically — empowered with nothing but standard tools of data analysis, such as PCA, filtering and cross-correlation analysis. Quite surprisingly we have found that even during quiet "no significant events" periods this standard set of statistical tools allows the "phasor-detective" to extract from the data important hidden anomalies, such as problematic control loops at loads and wind farms, and mildly malfunctioning assets, such as transformers and generators.**

I. Introduction

Data-driven techniques in power systems have at least fifty years of history, starting with static state estimations developed by Schweppe and co-authors [1]–[3], then transitioning to dynamic state estimation analysis and applications, see e.g. [4]–[6] and references therein, and most recently discussed under the umbrella of "big data" as the most significant enabler of power system operations, security and resiliency in the future [7], [8]. Many specific questions and approaches, including but not limited to modes of oscillations analysis of stability and detection [6], [9]–[11], dimension reduction for faster processing and analysis [12], early event detection [13], missing data recovery [14], identification of cyber attacks [15] and real time (online) event detection [16] are among the most recent advances.

On the methodology side data-driven methods developed in other engineering disciplines have been adopted, modified and used for many (e.g. aforementioned) power system applications. Principal component analysis [12]–[16], auto-correlation analysis of memory effects [17], and linear model driven spectral analysis of the dynamic state matrix [5], [6], [9]–[11], [18]–[20] are arguably the most popular data-driven methods currently in use in the power system research.

Even though the sophistication level of the methods already used in power system applications is impressive, coherence and understanding of the potential of new generation of the big data methods, driven during the last decade largely through heavy investment of IT industry, is still lacking [21]. We anticipate that many of the most modern methods, especially Deep Learning (DL) and related techniques [22]–[26], will impact the power-system operation-room reality in an even more significant ways. However, one problem with applications of the novel methods of DL and alike in sciences and engineering is that they are application agnostic/generic — very effective for many business cases, but lacking "explainability", i.e. intuitive physical/engineering explanations. This significant handicap of the most advanced and recent ML & AI methods slows down development of related applications in power systems. Indeed, power system practitioners would generally not consider as practical any new methods lacking "power systems informed" explanations.

This manuscript takes a step towards closing the gap between the rich variety of methods already developed and utilized in power systems and yet to be unleashed power of the upcoming Big Data revolution. Specifically, we start walking towards exciting sophistication of DL slowly, from the well-established and intuitive trenches of practical system engineering. We develop in this manuscript a pragmatic "phasor-detective" approach to analysis of the streaming Phasor Measurement Unit (PMU) data which allows to extract and interpret spatial and temporal correlations in a computationally light fashion and without making any constraining assumptions about origin of the correlations.

Our Contribution

We analyze synchronized historical PMU data recorded at ≈ 200 most significant locations of a US Independent System Operator (ISO) over the course of two years. At each PMU location the data includes complex current and voltage recorded with a millisecond resolution. Given geo-spatial locations of the PMU, but no information about the grid characteristics and layout, we pose the following principal questions: What can we possibly reconstruct from the data stream about the system current ambient behavior? To answer the question we utilize available statistical tools. In relation to preliminary data processing we apply to the raw signal two filtering techniques: sliding time horizon and Fourier analysis pre-processing. This allows to provide robust identification of the "quiet" periods and also prepare data for subsequent statistical analysis by means of Principal Component Analysis (PCA) and Auto-Correlation Analysis (ACA). We show that the two complementary tools, applied to the raw as well as to pre-processed signal, allow to separate scales and also provide compressed, thus easy to visualize, descriptors for online tracking of current state of the grid in a much broader way than what the current Energy Management System actually uses. PCA provides a robust set of indicators which record slow/adiabatic changes on the scale of seconds to

978-1-5386-4722-6/19/$31.00 ©2019 IEEE

ten of seconds and slower. We have observed that only a very few principal modes are significant at any moment of time, even though these modes may be different for voltage amplitude, phase difference and frequency (the three main characteristics) we track. The results do not change when the PCA is applied to the filtered signal, consistently with the fact that PCA averages over time but do not catch different-time correlations. ACA is the tool used to analyze the latter, in particular identifying significant, persistent correlations, missed by PCA, at shorter time scales — subseconds-to-seconds. Following ACA curves at different spatial locations we were able to identify nodes where correlations do not decay with time showing significant memory-effects. Remarkably, these nodes with significant memory cluster geographically. We observe two areas in the grid which show especially strong sustainable temporal correlations. We then proceed with ACA analysis of the Fourier-filtered signal. This helps us to identify and localize different harmonics. In particular, we observe (for a particular quiet period) emergence of significant oscillations in the 4-6Hz range at a small number of nodes. Interestingly, nodes with significant sustainable oscillations are either wind farms, big aggregated loads or mid-size generators. We conjecture that the sustainable oscillations are indicators of malfunction at these critical elements of the grid. We also observe that sustainable oscillations, seen clearly through emergence of a residue in the ACA analysis of the raw signal, disappear when applied to the Fourier-filtered signal (cutting off the 4-6Hz oscillations). Finally, analyzing spatial cross-correlations (of the residue) we were able to identify group of nodes with significant inter-dependency.

Material in the manuscript is organized as follows. Logic and main steps of the "phasor-detective" approach are described in Section II. Section III discusses the time series data we are working with and it also introduces averaging and filtering approaches we apply to the data. Covariance matrix and PCA analysis are discussed in Section IV. Auto-correlation and cross-correlation approaches, extracting information of the data temporal correlations/memory, are developed in Section V. Section VI is reserved for conclusions.

II. Logic and Steps of the Detective Approach

In this work we report on data streams from over 200 PMUs operated by an ISO and spanning a period over one year long. As is standard, the data stream from each PMU includes frequency, (complex) current and voltage reported 30 times per second. Using this data one can obtain real-time estimates of complex power at each location. Working with a data set this large (on the order of 28 TB) presents some obvious challenges; additionally there are specific artifacts that can arise in the data. For example, not all PMUs are always reporting, and occasionally some PMUs exhibit what appears to be errant behavior.

Our work has centered on performing statistical analysis aimed at inferring "structure" in the underlying transmission system as well as identifying complex behaviors, such as resonance and oscillations. In this manuscript we focus specifically on identification and characterization of "quiet" periods, also known as ambient conditions periods. Oversimplifying (see related discussion below) we consider a period *quiet* if fluctuations around the mean (e.g. characterized in terms of the standard deviations) are smaller than a reasonable pre-defined threshold. This focus on the quiet/ambient periods is motivated by the following considerations:

- The development of a strong understanding of quiet periods and (in particular) efficient online algorithms for recognition of such periods is a necessary step prior to studying less-quiet or even anomalous regimes, for otherwise we risk significant misinterpretation, i.e. errors in online detection of anomalies.
- As will be seen in this paper, the quiet regimes display informative patterns and correlations, all (slowly) time-evolving. Identifying such features is important with regards to:
  - Developing fast and reliable identification techniques.
  - Uncovering hidden malfunction of assets thus providing significant contribution towards forecasting most probable (and destined to occur) failures.
- The richness of correlations observed in the quiet regime, in fact, suggests that separation of what is normal/quiet from what is anomalous/atypical will be challenging. Even though we observe that quiet regimes dominate, relatively abrupt jumps of moderate size (i.e. jumps exceeding tracked standard deviation by factor of two or three) are rather frequent although overall they account for a relatively short fraction of the stream. As a result, it is rather difficult to find sufficiently long entirely quiet periods in the available data.
- Clearly, understanding quiet vs. volatile behavior will be helpful toward building predictive models for better optimization, control and planning.
- "Cyber-physical" attacks on power systems are a venue where fast and effective learning of (changing) stochastics may prove useful in identifying attacks. See [27].

The methodology adopted in this manuscript to identify the quiet periods is explained in Section III-A.

Notation: for $x \in \mathbb{C}$, $\bar{x}$ denotes its complex conjugate. For a complex matrix A, A^{H} denotes its Hermitian transpose.

III. Description and Averaging of the Time Series

The available data encompasses the period from January 1st 2013 to March 21st 2014. Each of the $N = 240$ PMUs records the following measurements 30 times per second: time of the measurement (GPS tagged), bus ID, voltage amplitude, voltage phase angle, current magnitude, current angle, and frequency. Additionally, the 2-dimensional coordinates of the PMU locations is also available, together with their corresponding nominal voltages. We note that PMUs *report* 30 times/second, but they *sample* at a far higher rate and perform filtering (e.g. anti-aliasing) before reporting.

We denote a generic scalar or complex measurement (e.g. complex voltage) at PMU location i, at time t by $m_i(t)$. The

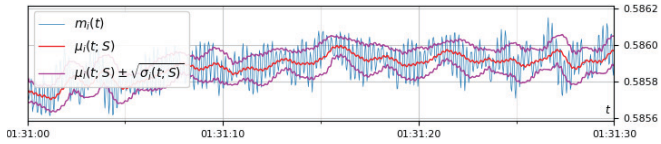

Fig. 1. Raw data $m_i(t)$ during 30 seconds of voltage magnitude (in blue) at a particular bus together with the average (in red) and a band of one standard deviation (in purple), using $S = 30$.

parameter t will be used to refer to the discrete time sequence with each temporal data point separated from preceding one by the same duration Δ ($1/30^{th}$ of a second in our case).

Typical pre-processing steps in the statistical analysis of data (especially with the goal of analyzing correlations) involve modifications through de-trending, offsetting (subtraction) of moving average, and normalizations. We will apply such techniques below. Specific details are provided next.

A. Averaging over Sliding Time Horizon & Quiet Periods

Consider a number S of measurements, this value corresponds to the memory budget. We define N dimensional vectors of means and variances (averaged over the (last) sliding time horizon of duration S) and the re-scaled zero-mean data vector, as follows, $\forall t$:

$$\mu_i(t;S) = \frac{1}{S}\sum_{\tau=t-S+1}^{t} m_i(\tau), \tag{1}$$

$$\sigma_i(t;S) = \frac{1}{S}\sum_{\tau=t-S+1}^{t} (m_i(\tau)-\mu_i(t;S))\,\overline{(m_i(\tau)-\mu_i(t;S))}, \tag{2}$$

$$\hat{m}_i(t;S) = \frac{m_i(t)-\mu_i(t-1;S)}{\sqrt{\sigma_i(t-1;S)}}. \tag{3}$$

Numerically, we found that a reasonable choice for S, that allows to separate power electronics (milliseconds) and electromechanical (seconds) time scales, is the number of readings in 1 second, that is $S = 30$. See Fig. 1.

Given a reference time t and a length parameter Q consider the $N \times Q$ matrix of normalized measurements

$$M(t;S;Q) = [\ \hat{m}_i(\tau;S)\ |\ \forall i,\ \tau \in \{t-Q+1,\ldots,t\}\ ],$$

corresponding to the last Q measurements before time t for all buses. We define the period $(t-Q,t]$ as *quiet* if the absolute value of all entries of the matrix $M(t;S;Q)$ is below some preset threshold. The reason behind this definition is that sudden jumps in the data appear as large values in the normalized time series, whereas normalized values close to zero mean that the data is behaving in a steady way. Effectively, a quiet period is an interval of time where all sensor-reported data behave in a stationary way. Moreover it is relatively cheap to compute $M(t;S;Q)$ from the stream data, as we just have to keep track of the last S measurements for each sensor, and the matrix of the last Q normalized measurements.

Again, in our analysis, we used $S = 30$ as the length of the sliding time horizon, and we consider quiet periods spanning 15 minutes. Over a selection of five different days across the database, we compute we compute the matrix $M(t;S;Q)$ and record its maximum absolute value when t spans over the complete day. Just in few cases the maximum was below 10 units, we selected 2-3 intervals between these recorded cases.

B. Fourier Filter

We have also applied Fourier filtering on the time series $m_i(\cdot)$. Let $\mathcal{F}[m_i(\cdot)] \in \mathbb{C}^Q$ be the discrete Fourier transform of $m_i(\cdot)$, where $\mathcal{F}[m_i(\cdot)]_k$ is the amplitude corresponding to frequency ω_k, $k \in \{0,\ldots,Q-1\}$. Since the frequency of the readings is 30Hz, we can represent the frequency domain as Q equidistant points between -15 and 15Hz.

We obtain a filtered time series by suppressing the high frequency components. Let $\mathbf{1}_{[-\lambda,\lambda]} \in \{0,1\}^Q$ be the indicator function of the set of frequencies in the interval $[-\lambda,\lambda]$. The filtered series is $\mathcal{F}^{-1}\left[\mathbf{1}_{[-\lambda,\lambda]} \odot \mathcal{F}[m_i(\cdot)]\right]$, where $\mathcal{F}^{-1}$ is the inverse Fourier transform and $\odot$ is the component-wise product. We denote this new series by $f_i(t;\lambda)$, $t \in \{0,\ldots,Q-1\}$.

Fig. 2a shows the amplitude of the Fourier transform of the complex voltage of a bus for an interval of 15 minutes. Fig. 2b-c show the magnitude of the signal and the filtered series when frequencies larger than 5 and 4Hz have been suppressed, respectively. Similarly, Fig. 2d-e show the angle of the signal and the filtered series when frequencies larger than 5 and 4Hz have been suppressed, respectively.

Analogously to (1)-(3), we compute (averaging over a sliding time horizon) the re-scaled zero-mean data vector for a real-valued series $f_i(\cdot;\lambda)$ instead of $m_i(\cdot)$, we call it $\hat{f}_i(\cdot;\lambda;S)$.

IV. Covariance Matrix & PCA analysis

Let $T \leq Q$ be the number of measurements for a period of length greater than the parameter S. The $N \times N$ covariance (correlation) matrix of a complex-valued signal, which is based on the last T measurements, equals, at time $T \leq t \leq Q$:

$$\Sigma_0(t;T;m^*(\cdot)) = \left[\ \frac{1}{T}\sum_{\tau=t-T+1}^{t} m_i^*(\tau)\overline{m_j^*(\tau)}\ \middle|\ \forall i,j\ \right], \tag{4}$$

where $m^*(\cdot)$ could be replaced by $\hat{m}(\cdot;S)$ or $\hat{f}(\cdot;\lambda;S)$. [Note that $\Sigma_0(t;T;\hat{m}(\cdot;S)) = \frac{1}{T}M(t;S;T)M(t;S;T)^{\mathsf{H}}$.]

We aim to choose a (fixed) value T that is sufficiently large so that the covariance is weakly dependent on T, but also not too large to keep memory as small as possible (with an eye on streaming applications). Empirical experiments show that with T corresponding to the number of measurements in three minutes we obtain stable covariance matrices when t varies.

In what follows we will drop the inputs T and $m^*(\cdot)$ from $\Sigma_0(t;T;m^*(\cdot))$ since they become fixed in the following analysis. We perform an eigen-decomposition on the correlation matrices as $\Sigma_0(t) = \sum_{k=1}^{N} \lambda_k(t)\xi_k(t)\xi_k(t)^{\mathsf{H}}$ and track the results as function of t, where $\xi_k(t) \in \mathbb{C}^N$ and $\lambda_k(t) \in \mathbb{R}$, $k \in \{1,\ldots,N\}$, are the orthonormal eigenvectors and corresponding eigenvalues (in decreasing order) of $\Sigma_0(t)$, respectively.

In our tests, $\Sigma_0(t)$ is (numerically) rank-deficient, justifying the PCA approach. PCA may be considered exact, if $K < N$ principal components are tracked or approximated. We approximate $\Sigma_0(t)$ replacing $N = 240$, by $K = 10$.

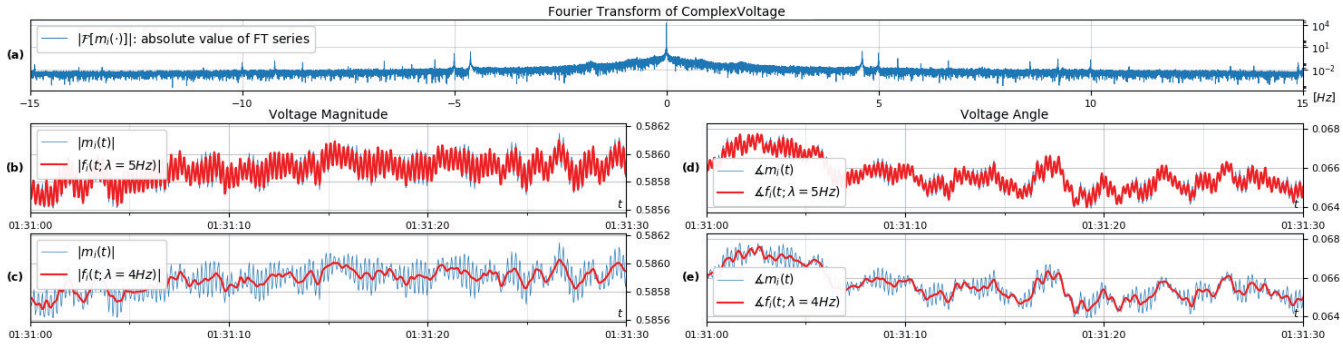

Fig. 2. **(a)** Absolute value (in log-scale) of Fourier transform of the complex voltage time series for a particular bus. Raw voltage magnitude and filtered series at **(b)** $\lambda = 5$Hz, and **(c)** $\lambda = 4$Hz. Raw voltage angle (in radians) and filtered series at **(d)** $\lambda = 5$Hz, and **(e)** $\lambda = 4$Hz.

The results of the PCA analysis are presented in Fig. 3 which are screenshots of the movies available in [28]. The movies show at time t for each of the three variables (frequency, voltage angle, and voltage magnitude) five indicators:

- normalized vector $m^*(t)$: for each sensor, we plot at its geographical position the value of the normalized data using different colors for different values (col. 1);
- first 40 eigenvalues of $\Sigma_0(t)$, in decreasing order (col. 2);
- largest 3 eigenvalues $\lambda_k(\cdot)$ for the last minute before t (cols. 3 to 5, upper section);
- eigenvectors $\xi_k(t)$ at time t, vector's component values are plotted geographically, in blue values close to -1 and in red values close to 1 (cols. 3 to 5, bottom section);
- auto-correlation function for a selected bus (col. 6 in the movie — omitted in Fig. 3), see Section V-A.

We also studied how the PCA changes if the received data is sparse, that is, from all the measurements made by the PMUs just a fraction of them are used in the analysis. As we mentioned above, PMUs report data at 30Hz, we performed PCA analyses considering only 6 readings per second and dropping the rest. The same spectrum of eigenvalues is obtained, though the later analysis has higher eigenvalues. Moreover, the 3 eigenvalue functions and eigenvectors are similar to Fig. 3.

V. Accounting for Temporal Correlations

Consider the delayed covariance matrix generalizing (4), $\forall \Delta \geq 0,\ \forall t \geq T + \Delta$:

$$\Sigma_\Delta(t;T;m^*(\cdot)) = \left[\frac{1}{T} \sum_{\tau=t-T+1}^{t} m_i^*(\tau)\overline{m_j^*(\tau-\Delta)} \;\middle|\; \forall i,j \right]. \quad (5)$$

We are interested to study how $\Sigma_\Delta(t;T;m^*(\cdot))$ changes with Δ increase and then track evolution with t.

However, evolution of the spectrum (in the two-dimensional (Δ,t) space) is challenging. Instead, we study two surrogate objects, introduced in the following two subsections, which are easier to visualize. Again, since T and $m^*(\cdot)$ are fixed, we will omit them as input of the Σ_Δ function.

A. Auto-Correlation Functions

The normalized PMU's auto-correlation functions are:

$$\forall \text{ node } i: \qquad \mathcal{A}_i(\Delta;t) = [\Sigma_\Delta(t)]_{ii} / |[\Sigma_0(t)]_{ii}|. \quad (6)$$

This object is of interest because of the following two reasons:

- Dependence of the auto-correlation function on Δ indicates whether fluctuations around the mean at a particular node decay or not with time. Stated differently this analysis tests if there are significant memory effects or if memory is lost.
- It accounts for the part of the measurement matrix which is ignored in the PCA analysis, as discussed above –that is, it accounts for the temporal matrix W, that comes from the singular-value decomposition $M(t;S;T) = UDW^\top$.

We show dependence of the auto-correlations on time, i.e. dependence of $\mathcal{A}_i(\Delta;t)$ on t. The movies are available in [28] and example snapshots are given on the two left plots of Fig. 4a, which shows two auto-correlation functions $\mathcal{A}_i(\cdot;t)$ for selected nodes at a specific time t. These nodes were chosen between the ones with largest auto-correlation amplitude.

We also apply the auto-correlation analysis to the filtered signal. The results of filtering the row signal, correspondent to the Fig. 4a, at 5 and 4Hz (cutting off higher harmonics) are shown in Fig. 4b-c. We observe that if the signal is filtered at a sufficiently low frequency (4Hz) in this case a significant level of oscillations, seen otherwise in Fig. 4a-b, disappear.

In order to measure this amplitude and formalize the concept of the *residue* one studies,

$$\forall i: \quad \rho_i([\Delta_{min},\Delta_{max}];t) = \frac{1}{\Delta_{\max} - \Delta_{min}} \sum_{\Delta=\Delta_{min}}^{\Delta_{max}-1} |\mathcal{A}_i(\Delta;t)|, \quad (7)$$

where Δ_{min} was set to $1s$ (in order to ignore the firsts values of $\mathcal{A}_i(\cdot;t)$) and Δ_{max} was set to $1min = 60s$ in our tests. (We have verified empirically that with this value of Δ_{max} the results are robust with $O(1)$ changes in Δ_{max}.)

Third column of Fig. 4 geographically shows the residue for each node when different normalized time series are considered. As seen in the movies of the residue maps, residue values at a number of special nodes do not decay with time. Moreover we observe that nodes with a significant residue cluster, specifically, there are two sets of nodes that repeat their high values across different days, months and time of the day. Those groups are also shown in Fig. 4b: in the bottom left part of the map and the upper right zone (closer to the center); they might show up together or by separate.

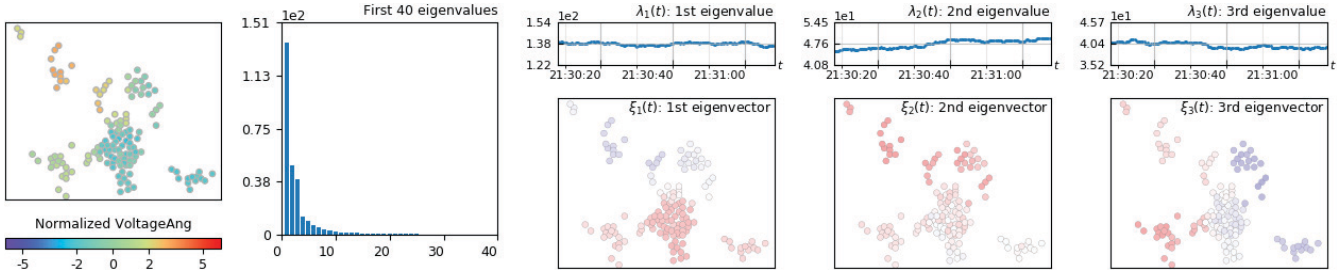

Fig. 3. Movie screenshot of a selected quiet period at a specific moment t, showing the voltage phase angle PCA analysis. Here, $m^*(\cdot) = \hat{m}(\cdot, S = 1\text{sec})$ and $T = 3\text{min}$. Columns from left to right: normalized phase angle (3) at all buses, eigenvalues of $\Sigma_0(t)$, first, second, and third largest eigenvalues during the last minute (above) together with their corresponding normalized eigenvectors (below).

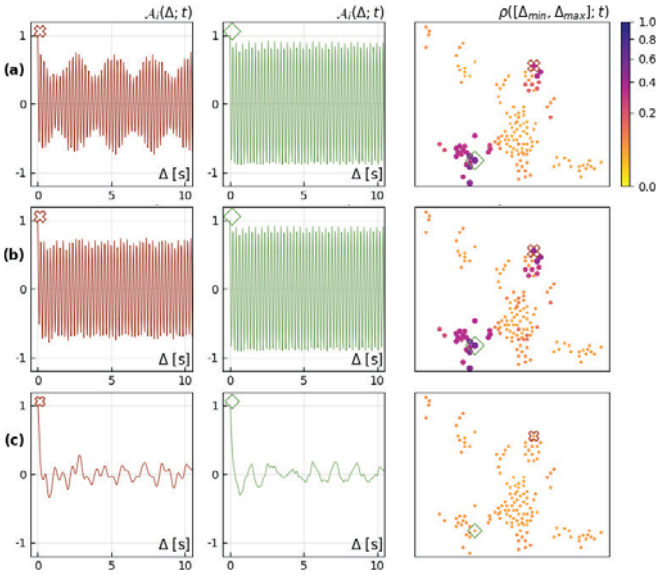

Fig. 4. Movie screenshot of the auto-correlation functions for frequency at two selected buses at a specific moment t. Here, $S = 1\text{sec}$ and $T = 60\text{sec}$. Different rows show the result using different time series: **(a)** $\hat{m}(\cdot; S)$, **(b)** $\hat{f}(\cdot; \lambda = 5\text{Hz}; S)$, and **(c)** $\hat{f}(\cdot; \lambda = 4\text{Hz}; S)$. Last column shows the residue for all nodes, geometric figures (cross and diamond) correspond to the position of the buses whose auto-correlation function are shown in the first two columns.

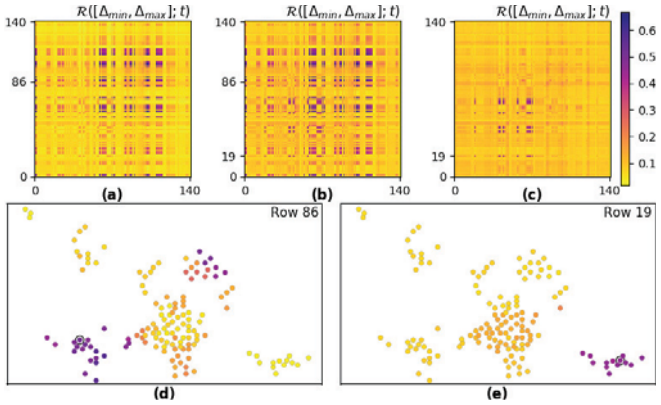

Fig. 5. Cross-Correlation matrices of 141 buses using **(a)** $\hat{m}(\cdot; S)$, **(b)** $\hat{f}(\cdot; \lambda = 5\text{Hz}; S)$, and **(c)** $\hat{f}(\cdot; \lambda = 4\text{Hz}; S)$. Here, $T = 60\text{sec}$, $S = 1\text{sec}$, $\Delta_{min} = 1\text{sec}$, and $\Delta_{max} = 60\text{sec}$. Geographical map with values of **(d)** row 86 of (a) and **(e)** row 19 of (c). Buses corresponding to each row have an extra outer-circle.

Observation that sustainable correlations stay for sufficiently long period of time suggest to analyze spatio-temporal features of the sustainable correlations via the Cross-Correlation Residue described in the next Subsection.

B. Cross-Correlation Residue (CCR)

The cross-correlation version of (7) is, $\forall i, j$:

$$\mathcal{R}_{ij}([\Delta_{min}, \Delta_{max}]; t) = \frac{1}{\Delta_{max} - \Delta_{min}} \sum_{\Delta=\Delta_{min}}^{\Delta_{max}-1} |[\Sigma_\Delta(t)]_{ij}|, \quad (8)$$

where we drop normalization to avoid singularities associated with signals at different nodes which are not correlated. To visualize the CCR (8), we plot the matrix $[\mathcal{R}_{ij}([\Delta_{min}, \Delta_{max}]; t)]_{ij}$ showing in darker colors the higher values, see Fig. 5a-c.

The three plots in Fig. 5a-c show the CCR matrix at the same time using different time series, and depending on this, different patterns appear. In Fig. 5a 27 buses have in its corresponding row some component with a value higher than 0.4, whereas in Fig. 5c only 13 buses and different from the ones in Fig. 5a. Fig. 5b shows 40 buses with values higher than 0.4, the same as Fig. 5a and Fig. 5c together.

Fig. 5d-e shows geographically the value that each bus has in row 86 of Fig. 5a and row 19 of Fig. 5c. We can see from these plots that high cross-correlation between buses is strongly connected with geographical closeness. We also note that components that have high auto-correlation in Fig. 4 have also high cross-correlation between themselves.

VI. Conclusions

In this Section we summarize the results it has helped to discover analyzing selected "quiet" periods in the PMU data recorded over the period of two years by an ISO.

• Averaged over time PMU signal takes care of noise in the data and shows interesting spatial correlations. Correlations are different for different objects of interest (frequency, phase and voltage) and also different for different quiet periods. Matrix of correlations is sparse, also revealing that number of the high-intensity contributions to the correlations is small. (Note that the statement of sparseness is consistent with previous studies of the measurement matrix, see e.g. [16].) Each of the contributions characterizes a mode localized on a relatively few

nodes (PMU positions) within the system. Principal modes, computed over ambient conditions, are almost frozen in time, however responding fast to any significant perturbation, thus suggesting them as efficient features/indicators for changes.

• Fourier-analysis of the signal reveals interesting spatial patterns. Extracting modes in the 4-6Hz range one observes significant contributions from only few PMU nodes. The spatial patterns are different for the thee different characteristics and they are also different for different periods. Nodes showing large 4-6Hz contributions were identified as aggregated loads, mid-size generators and large wind-farms.

• In general the signals are long-correlated in time. However the memory effects becomes significantly less prominent when Fourier filtering at 4Hz, cutting higher frequencies, is applied. Nodes with significant auto-correlations (memory) are relatively few and the respective spatial pattern is adiabatic/frozen (changing slowly during the analyzed periods). Like other adiabatic patterns mentioned above, the pattern changes from one quiet period to another and we also observed different patterns for different characteristics (frequency, phase, voltage).

• Analyzing time-delayed cross-correlations between different nodes we observe that, like in the case of the auto-correlations, correlations between some nodes have long memory. Nodes which mutual inference shows a long memory form a sparse pattern. These patterns, like others described above are adiabatic and evolving from characteristic to characteristic and from one quiet to another quiet period.

• Finally, these algorithms can be applied in an online fashion and provide a visual tool to supervise the network functionality (as shown in [28]), our simulations showed almost no changes between using (1)-(3) and a moving average version of them. It suffices to keep track of the $T+\Delta$ last measurements to update the correlation matrices Σ_Δ. The most costly algorithm is the eigen-decomposition, that has complexity $O(N^3)$, however, this can be easily done every few seconds, for the data we have experimented with (available at $N = 240$ locations).

ACKNOWLEDGMENTS

The work at LANL was carried out under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy under Contract No. DE-AC52-06NA25396. The work was partially supported by DOE/OE/GMLC and LANL/LDRD/CNLS projects.

REFERENCES

[1] F. Schweppe and J. Wildes, "Power System Static-State Estimation, Part I: Exact Model," *IEEE Trans. on Power Apparatus and Systems*, vol. PAS-89, no. 1, pp. 120–125, Jan 1970.

[2] F. Schweppe and D. Rom, "Power System Static-State Estimation, Part II: Approximate Model," *IEEE Trans. on Power Apparatus and Systems*, vol. PAS-89, no. 1, pp. 125–130, Jan 1970.

[3] F. Schweppe, "Power System Static-State Estimation, Part III: Implementation," *IEEE Trans. on Power Apparatus and Systems*, vol. PAS-89, no. 1, pp. 130–135, Jan 1970.

[4] J. Hauer, C. Demeure, and L. Scharf, "Initial results in Prony analysis of power system response signals," *IEEE Trans. on Power Systems*, vol. 5, no. 1, pp. 80–89, Feb 1990.

[5] G. Liu, J. Quintero, and V. Venkatasubramanian, "Oscillation monitoring system based on wide area synchrophasors in power systems," in *2007 iREP Symposium - Bulk Power System Dynamics and Control - VII. Revitalizing Operational Reliability*, Aug 2007, pp. 1–13.

[6] D. Trudnowski, "Properties of the Dominant Inter-Area Modes in the WECC Interconnect," https://www.wecc.biz/Reliability/WECCmodesPaper130113Trudnowski.pdf, 2012, accessed: 2018-09-11.

[7] M. Kezunovic, L. Xie, and S. Grijalva, "The role of big data in improving power system operation and protection," in *2013 IREP Symposium Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid*, Aug 2013, pp. 1–9.

[8] R. Arghandeh and Y. Zhou, *Big Data Application in Power Systems*. Elsevier, 2018.

[9] "Reliability Guideline Forced Oscillation Monitoring & Mitigation," https://www.nerc.com/pa/RAPA/rg/ReliabilityGuidelines/Reliability_Guideline_-_Forced_Oscillations_-_2017.pdf, 2017, accessed: 2018-09-11.

[10] "Power System Oscillatory Behaviors: Sources, Characteristics, & Analyses," https://www.pnnl.gov/main/publications/external/technical_reports/PNNL-26375.pdf, 2017, accessed: 2018-09-11.

[11] "Report of WECC Joint Synchronized Information Subcommittee on Modes of Inter-Area Power Oscillations in Western Interconnection," https://www.wecc.biz/Reliability/WECC%20JSIS%20Modes%20of%20Inter-Area%20Oscillations-2013-12-REV1.1.pdf, 2013, accessed: 2018-09-11.

[12] N. Dahal, R. King, and V. Madani, "Online dimension reduction of synchrophasor data," *Transmission and Distribution Conference and Exposition (T&D)*, pp. 1–7, May 2012.

[13] L. Xie, Y. Chen, and P. Kumar, "Dimensionality Reduction of Synchrophasor Data for Early Event Detection: Linearized Analysis," *IEEE Trans. on Power Systems*, vol. 29, no. 6, pp. 2784–2794, 2014.

[14] P. Gao, M. Wang, S. Ghiocel, J. Chow, B. Fardanesh, and G. Stefopoulos, "Missing Data Recovery by Exploiting Low-Dimensionality in Power System Synchrophasor Measurements," *IEEE Trans. on Power Systems*, vol. 31, no. 2, pp. 1006–1013, March 2016.

[15] P. Gao, M. Wang, J. Chow, S. Ghiocel, B. Fardanesh, G. Stefopoulos, and M. Razanousky, "Identification of Successive Unobservable Cyber Data Attacks in Power Systems Through Matrix Decomposition," *IEEE Trans. on Signal Processing*, vol. 64, no. 21, pp. 5557–5570, Nov 2016.

[16] W. Li, M. Wang, and J. Chow, "Real-Time Event Identification Through Low-Dimensional Subspace Characterization of High-Dimensional Synchrophasor Data," *IEEE Trans. on Power Systems*, vol. 33, no. 5, pp. 4937–4947, Sept 2018.

[17] S. Roy and B. Lesieutre, "Frequency Band Decomposition of a Dynamic Persistence Measure Using Ambient Synchrophasor Data," in *2018 Power Systems Computation Conference (PSCC)*, June 2018, pp. 1–7.

[18] C. Chang and Z. Li, "Recursive stochastic subspace identification for structural parameter estimation," in *Proc. SPIE*, vol. 7292, Mar 2009.

[19] S. Sarmadi and V. Venkatasubramanian, "Inter-Area Resonance in Power Systems From Forced Oscillations," *IEEE Trans. on Power Systems*, vol. 31, no. 1, pp. 378–386, Jan 2016.

[20] A. Lokhov, M. Vuffray, D. Shemetov, D. Deka, and M. Chertkov, "Online Learning of Power Transmission Dynamics," *PSCC 2018, arXiv:1710.10021*, 2017.

[21] "Department of energy: Big data analysis of synchrophasor data," https://grantbulletin.research.uiowa.edu/netl-big-data-analysis-synchrophasor-data, 2018.

[22] Y. Bengio, "Learning Deep Architectures for AI," *Found. & Trends Machine Learning*, vol. 2, no. 1, pp. 1–127, Jan. 2009.

[23] I. Arel, D. Rose, and T. Karnowski, "Deep Machine Learning - A New Frontier in Artificial Intelligence Research [Research Frontier]," *IEEE Computational Intelligence Mag.*, vol. 5, no. 4, pp. 13–18, Nov 2010.

[24] J. Schmidhuber, "Deep Learning in Neural Networks," *Neural Networks*, vol. 61, no. C, pp. 85–117, Jan. 2015.

[25] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[26] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT Press, 2016.

[27] D. Bienstock and M. Escobar, "Stochastic Defense Against Complex Grid Attacks," *arXiv:1807.06707*, 2018.

[28] D. Bienstock, M. Chertkov, and M. Escobar, "Learning from ISO-scale PMU data stream," 2018. [Online]. Available: http://www.columbia.edu/~dano/research/pgrid/pmus/index.html